

Supplementary Material of “FedEvi: Improving Federated Medical Image Segmentation via Evidential Weight Aggregation”

Jiayi Chen^{1*}, Benteng Ma^{2*}, Hengfei Cui¹, and Yong Xia^{1,3,4}(✉)

¹ National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science and Engineering, Northwestern Polytechnical University, Xi’an 710072, China

yxia@nwpu.edu.cn

² Hong Kong University of Science and Technology, Hong Kong SAR, China

³ Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China

⁴ Ningbo Institute of Northwestern Polytechnical University, Ningbo 315048, China

Algorithm A1 Algorithm of the proposed FedEvi

Input: global model θ , local models $\{\theta_k\}_{k=1}^K$, local datasets $\{\mathcal{D}_k\}_{k=1}^K$, aggregation weights $\{\beta_k\}_{k=1}^K$, federated rounds R , local training epochs E

Output: aggregated global model θ^R

- 1: **for** $r = 1$ **to** R **do**
 - 2: **Server:**
 - 3: Distribute θ^r to K local clients, updating θ_k to $\theta_k \leftarrow \theta^r$
 - 4: **Client:**
 - 5: $\theta_k^r \leftarrow \text{LocalTraining}(\mathcal{D}_k^{\text{train}}, \theta_k, E)$
 - 6: Upload $\{\theta_k^r\}_{k=1}^K$ to the server.
 - 7: **Server:**
 - 8: Construct the surrogate model $\hat{\theta}^r$ as $\hat{\theta}^r = \sum_{k=1}^K \beta_k^{r-1} \theta_k^r$.
 - 9: Distribute $\hat{\theta}^r$ to K local clients.
 - 10: **Client:**
 - 11: Measure $G(\mathcal{D}_k^{\text{val}}, \hat{\theta}^r)$ and $R(\mathcal{D}_k^{\text{val}}, \theta_k^r)$ for each client using Eq. 4 and Eq. 5.
 - 12: Upload $\{G(\mathcal{D}_k^{\text{val}}, \hat{\theta}^r)\}_{k=1}^K$ and $\{R(\mathcal{D}_k^{\text{val}}, \theta_k^r)\}_{k=1}^K$ to the server.
 - 13: **Server:**
 - 14: Adjust $\{\beta_k^{r-1}\}_{k=1}^K$ to obtain $\{\beta_k^r\}_{k=1}^K$ based on Eq. 8.
 - 15: Aggregate $\{\theta_k^r\}_{k=1}^K$ to attain the global model as $\theta^{r+1} = \sum_{k=1}^K \beta_k^r \theta_k^r$.
 - 16: **end for**
-

Table A1. Data source and sample sizes (train/val/test) of multi-center datasets utilized in our study.

Dataset	Endoscopic Polyp Dataset	Prostate MRI Dataset	Retinal Fundus Dataset
Client 1	Kvasir [9]: 700/100/200	BIDMC [17]: 207/16/38	Drishiti-GS [30]: 71/10/20
Client 2	ETIS [29]: 138/19/39	BMC [3]: 276/38/70	RIM-ONE-r3 [7]: 113/15/31
Client 3	ColonDB [31]: 267/37/75	HK [17]: 121/11/26	REFUGE (Zeiss) [23]: 280/40/80
Client 4	ClinicDB [2]: 429/61/122	I2CVB [12]: 358/28/82	REFUGE (Canon) [23]: 280/40/80
Client 5	-	RUNMC [3]: 295/51/75	BinRushed [1]: 67/9/19
Client 6	-	UCL [17]: 138/13/24	Margabia [1]: 137/19/39

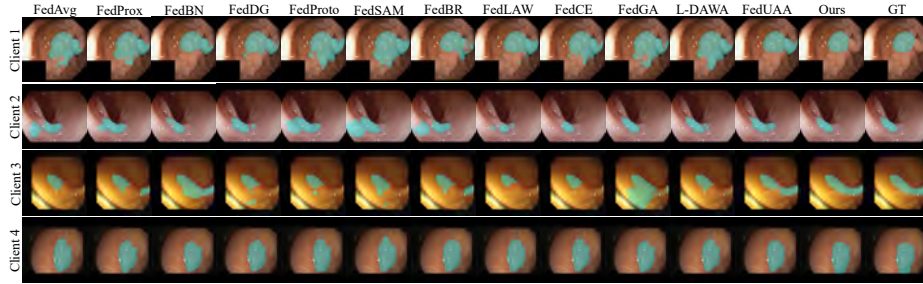


Fig. A1. Visualization of segmentation results obtained by 12 competing methods and our FedEvi on endoscopic polyp segmentation.

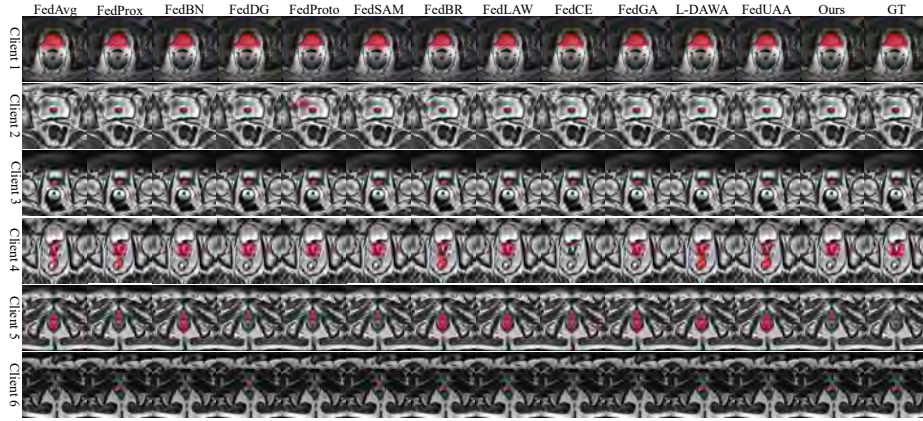


Fig. A2. Visualization of segmentation results obtained by 12 competing methods and our FedEvi on prostate MRI segmentation.

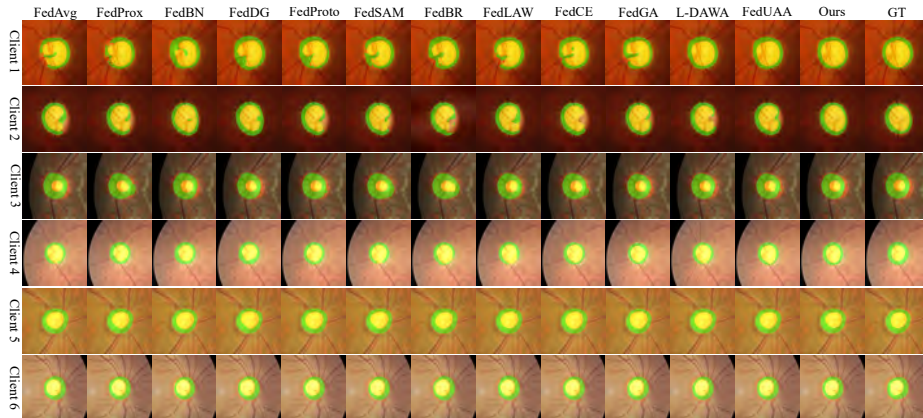


Fig. A3. Visualization of segmentation results obtained by 12 competing methods and our FedEvi on retinal fundus segmentation.